A Remark on Mordell's Conjecture

A REMARK ON MORDELL'S CONJECTURE.

By DAVID MUMFORD.*

It is somewhat surprising that the systematic evaluation of the heights of rational points on a curve and on its jacobian variety and particularly of their relation to each other should yield any new information. Nonetheless this appears to be the case and the result is described in this article. Although the main theorem is not even a special case of the very fascinating conjecture of Mordell, still it is an estimate that already reveals that rational points on curves of genus at least 2 are much harder to come by than on curves of genus 0 or 1. It is a quantitative limitation on the heights of such points which is well-known to be false in the case of genus 0 or 1. Incidentally, there is a good explanation why an estimate of this type can be obtained so cheaply, whereas Mordell's conjecture itself could not: namely, results obtained by our methods will more or less automatically apply to the analogous "function field" case [where the ground field is a function field in one variable over a finite field, rather than an algebraic number field]. And in this case, unless further restrictions are imposed, there are curves of any genus with an infinite number of rational points whose heights increase exactly at the rate which we will find.

Let k be an algebraic number field of finite degree over $\boldsymbol{Q}$. Let C be a non-singular projective curve over k of genus g at least 2. Mordell's conjecture asserts that the set of k-rational points on C is finite. Now suppose that a projective embedding of C is fixed, allowing us to talk of the heights, $ht(x)$, of k-rational points of x. Then my result is this:

THEOREM. *There are real constants a and b, $a > 0$, such that if the countable set of k-rational points of C is ordered by increasing height—call the points $x_1, x_2, \cdots$—then*

$$ht(x_i) \geqq e^{ai+b}.$$

Because of the well-known properties of heights, this result is not affected by changing the projective embedding of C. An example of the theorem is given by Fermat's curve:

Received February 25, 1965.

* This research was partially supported by the AMS 1964 Summer Institute in Algebraic Geometry and NSF-GP3512.

COROLLARY. *Let $(\alpha_i, \beta_i, \gamma_i)$ be an infinite set of distinct positive integral solutions of the equation*

$$X^n + Y^n = Z^n$$

such that $\alpha_i, \beta_i, \gamma_i$ have no common factors and such that $\{\gamma_i\}$ is an increasing sequence. Assume $n \geqq 4$. There are real constants a and b, $a > 0$, such that

$$\gamma_i \geqq e^{(e^{ai+b})}.$$

A final word: that the proof of the theorem appears in as natural and simple a form as it does is due to the collaboration of John Tate; that it appears in print, needless to say, is not.

1. The theory of heights. We fix an algebraic number field k, of finite degree over $\boldsymbol{Q}$. The main result of the "classical" Theory of Weil (cf. [1] and [4]) is the construction of a set of functions as follows:
Given: a scheme X, projective over k, and an element $\delta \in Pic(X)$.
Constuct: a real-valued function on the set of k-rational points X_k, written

$$h_\delta(x),\ x \in V_k$$

In fact, h_δ is not constructed precisely, but only the class of all functions, differing from one member of this class by a *bounded* function is constructed. This construction has the following properties (where $O(x)$ denotes a bounded function of x):

a) If $f: X \to Y$ is a k-morphism of schemes X and Y as above, and if $\delta \in Pic(Y)$, then

$$h_\delta(f(x)) = h_{f^*(\delta)}(x) + O(x)$$

b) If $\delta_1, \delta_2 \in Pic(X)$, for X as above, then

$$h_{\delta_1+\delta_2}(x) = h_{\delta_1}(x) + h_{\delta_2}(x) + O(x)$$

c) If D is an effective Cartier divisor on the projective scheme X, and if D defines the element $\delta \in Pic(X)$, then there is a real constant K such that

$$h_\delta(x) \geqq K, \text{ all } x \in X - \text{Support } (D).$$

d) If $\delta \in Pic(X)$ is ample, then for all constants K, the set of points $x \in X_k$ such that $h_\delta(x) \leqq K$ is finite.

The lack of a really definite height function is one of the most awkward aspects of this theory. In case X is assumed to be an abelian variety, this

defest has been remedied by Néron and Tate (cf. [2], [3], [4½]). The simplest way to state their result is this:

THEOREM. *Let X be an abelian variety, and let $\delta \in Pic(X)$. Then the class of functions h_δ on X contains a "quadratic" function on X, i.e., a function f satisfying the identity:*

$$f(x+y+z) - f(x+y) - f(x+z) - f(y+z) + f(x) + f(y) + f(z) - f(0) = 0.$$

One checks immediately that a real-valued bounded quadratic function is constant. Therefore, if we put the two requirements on h_δ that (1) it is quadratic, and (2) it is 0 at the identity point e, then we obtain a completely well-defined height function. Moreover, we get the important Corollary:

COROLLARY. 1) *If X is an abelian variety, and $\delta_1, \delta_2 \in Pic(X)$, then the normalized height functions on X satisfy:*

$$h_{\delta_1+\delta_2}(x) = h_{\delta_1}(x) + h_{\delta_2}(x), \text{ all } x \in X_k.$$

2) *If $f: X \to Y$ is any morphism of abelian varieties, and $\delta \in Pic(Y)$, then*

$$h_{f^*\delta}(x) = h_\delta(f(x)) - h_\delta(f(e)),$$

all $x \in X_k$. In particular, if f is a homomorphism (i.e., takes the identity to the identity), then

$$h_{f^*\delta}(x) = h_\delta(f(x)).$$

2. The set-up derived from a curve. We shall assume given a non-singular projective curve C, over k, with genus $g \geqq 1$. The purpose of this section is to give a thorough account of the auxiliary varieties associated to C, the canonical divisor classes that they carry, and their universal properties. For the sake of simplicity, we also assume that a base point $x_0 \in C_k$ has been chosen once and for all; and that all other schemes X occurring in the discussion have base points p_X. (The base points on abelian varieties will be assumed to be their identity points). - A general concept which is central to the discussion is the following:

Definition. Let X and Y be connected algebraic schemes over k. A *divisorial correspondence* on $X \times Y$ is an element $\delta \in Pic(X \times Y)$ which is 0 restricted to either of the subschemes $X \times \{p_Y\}$ of $\{p_X\} \times Y$.

First of all, let J be the connected component of the identity of the

Picard scheme of C: i.e., the so-called "Jacobian variety" of C. It is an abelian variety of dimension g. Moreover, J is characterized by the existence of a canonical divisorial correspondence

$$\delta_1 \in Pic(C \times J)$$

which has the universal mapping property (cf. [5] and [6]):

(*) For all connected algebraic schemes X, and all divisorial correspondences η on $C \times X$, there is a unique morphism $f: X \to J$ such that

$$(1_C \times f)^*(\delta_1) = \eta.$$

Secondly, on the non-singular surface $C \times C$ the Weil divisor

$$\Delta - C \times \{x_0\} - \{x_0\} \times C$$

defines an element $\Delta^* \in Pic(C \times C)$ which is clearly a divisorial correspondence. By the $UMP(*)$, there is a unique morphism

$$\phi: C \to J$$

such that $\Delta^* = (1_C \times \phi)^*(\delta_1)$.

Thirdly, let $\hat{J}$ be the connected component of the identity of the Picard scheme of J: i.e., the dual abelian variety. J is characterized by the existence of a canonical divisorial correspondence

$$\delta_2 \in Pic(J \times \hat{J})$$

which has the universal mapping property:

(**) For all connected algebraic schemes X, and all divisorial correspondences η on $J \times X$, there is a unique morphism $f: X \to \hat{J}$ such that

$$(1_J \times f)^*(\delta_2) = \eta.$$

Fourthly, the morphism ϕ dualizes to a morphism $\hat{\phi}: \hat{J} \to J$. Namely, apply the Universal mapping property (*) with $X = J$, $\eta = (\phi \times 1_{\hat{J}})^*(\delta_2)$. This means that we get a diagram:

(***)
$$\begin{array}{ccc} C \times \hat{J} & \xrightarrow{\phi \times 1_{\hat{J}}} & J \times \hat{J} \\ \downarrow{\scriptstyle 1_C \times \hat{\phi}} & & \\ C \times J & & \end{array}$$

such that δ_1 and δ_2 induce the same correspondence on $C \times \hat{J}$.

Fifthly, recall the general construction by which divisor classes η on abelian varieties X define homomorphisms from X to its dual $\hat{X}$. There are three maps from $X \times X$ to X—the group law μ and the two projections p_1 and p_2. Then one checks that for any $\eta \in Pic(X)$, the divisor class

$$\mu^*(\eta) - p_1^*(\eta) - p_2^*(\eta)$$

is a divisorial correspondence on $X \times X$. Therefore, by definition of $\hat{X}$, there is a unique morphism $f: X \to \hat{X}$ such that

$$\mu^*\eta - p_1^*\eta - p_2^*\eta = (1_X \times f)^* \begin{bmatrix} \text{canonical class} \\ \text{on } X \times \hat{X} \end{bmatrix}.$$

We will denote f by $\Lambda(\eta)$. Recall that Λ is itself a homomorphism: $\Lambda(\eta_1 \pm \eta_2) = \Lambda(\eta_1) \pm \Lambda(\eta_2)$. In terms of this definition, the central result concerning jacobians is the following (due to Weil [7]).

THEOREM. *$\exists$ an ample divisor Θ on J such that*

$$\hat{\phi} = -\Lambda(\Theta)^{-1}.$$

In fact, recall that Θ is nothing but the sum of the subset $\phi(C)$ in J with itself (with respect to the group law in J) $(g-1)$ times. For reference we write the meaning of this Theorem out as follows:

$$\begin{cases} \psi = -\hat{\phi}^{-1} \\ \text{class of } \underbrace{\mu^*\Theta - p_1^*\Theta - p_2^*\Theta}_{\text{call this } \theta} = (1_J \times \psi)^*(\delta_2). \end{cases}$$

The net result of all this is the following: suppose we identify J with $\hat{J}$ via the isomorphism ψ, or $\Lambda(\Theta)$. Then we have defined the canonical divisor classes:

On $C \times C$: Δ^*

On $C \times J$: δ_1

On $J \times J$: $\theta = \text{class of } \mu^*\Theta - p_1^*\Theta - p_2^*\Theta$
$= \delta_2$
via our identifications

These are related by the equations

(a) $\Delta^* = (1_C \times \phi)^*(\delta_1)$

(b) $\delta_1 = -(\phi \times 1_J)^*(\theta)$.

hence

(c) $\Delta^* = -(\phi \times \phi)^*(\theta)$.

Proof. (a) has been pointed out before, and (c) follows from (a) and (b). As for (b), first use the fact that $\Lambda(-\Theta) = -\Lambda(\Theta) = -\psi$. Therefore

$$-\theta = (1_J \times (-\psi))^*\delta_2 = (1_J \times \hat{\phi}^{-1})^*\delta_2.$$

Hence

$$-(1_J \times \hat{\phi})^*\theta = \delta_2$$

and finally:

$$\begin{aligned}(1_C \times \hat{\phi})^*\delta_1 &= (\phi \times 1_{\hat{J}})^*\delta_2 \qquad \text{(This is (***))} \\ &= -(\phi \times 1_{\hat{J}})^*(1_J \times \hat{\phi})^*\theta \\ &= -(\phi \times \hat{\phi})^*\theta \\ &= -(1_C \times \hat{\phi})^*(\phi \times 1_{\hat{J}})^*\theta.\end{aligned}$$

Since $1_C \times \hat{\phi}$ is an isomorphism, (b) follows. Q. E. D.

3. The basic estimates. Once again, we consider a curve C over a number field k, as above. Now we will use the maps obtained in § 2 to obtain properties of the height functions introduced in § 1. The most important height function is $h_\theta(x, y)$ defined for $x, y \in J_k$.

PROPOSITION 1. *$h_\theta(x, y)$ is a symmetric, bilinear form on $J_k \times J_k$. Moreover it is positive definite on J_k/mod torsion.*

Proof. Let $f_1 : J \to J \times J$ be the homomorphism mapping x to $x \times e$, and let f_2 map x to $e \times x$. Since θ is a divisorial correspondence, $f_1^*\theta = f_2^*\theta = 0$. Therefore

$$\begin{aligned}h_\theta(x, e) &= h_\theta(f_1(x)) = h_{f_1^*\theta}(x) = 0, \\ h_\theta(e, x) &= h_\theta(f_2(x)) = h_{f_2^*\theta}(x) = 0.\end{aligned}$$

But this means that h_θ is a quadratic function on the product of two groups which is 0 on both factors alone. It is easy to check that this implies that h_θ is bilinear.

Let $\xi : J \times J \to J \times J$ be the morphism mapping $x \times y$ to $y \times x$. Then clearly $\xi^*\theta = \theta$, hence

$$h_\theta(x, y) = h_\theta(\xi(y, x)) = h_{\xi^*\theta}(y, x) = h_\theta(y, x).$$

To evaluate $h_\theta(x, x)$, let $\Delta : J \to J \times J$ be the diagonal morphism, and let $\lambda_2 : J \to J$ be multiplication by 2. Then

$$\begin{aligned} h_\theta(x,x) &= h_\theta(\Delta(x)) \\ &= h_{\Delta^*\theta}(x) \\ &= h_{\Delta^*(\mu^*\Theta - p_1^*\Theta - p_2^*\Theta)}(x) \\ &= h_{\lambda_2^*\Theta}(x) - 2h_\Theta(x). \end{aligned}$$

since $\lambda_2 = \mu \circ \Delta$, $1_J = p_i \circ \Delta$. On the other hand, if D is any divisor on J, let D' be the divisor obtained by reflecting D in the origin. Then $\lambda_2^*(D)$ is in the same divisor class as $3D + D'$. Therefore,

$$\begin{aligned} h_\theta(x,x) &= h_\Theta(x) + h_{\Theta'}(x) \\ &= h_\Theta(x) + h_\Theta(-x) \end{aligned}$$

I claim that if this is not positive, then x must be a torsion point on J. Namely, assume $h_\theta(x,x) \leqq 0$. Then for all integers n,

$$\begin{aligned} h_\Theta(nx) + h_\Theta(-nx) &= h_\theta(nx, nx) \\ &= n^2 h_\theta(x,x) \\ &\leqq 0, \end{aligned}$$

hence either $h_\Theta(nx) \leqq 0$ or $h_\Theta(-nx) \leqq 0$. This means that if x is not a torsion point, there are an infinite number of distinct points x_i such that $h_\Theta(x_i) \leqq 0$. Since Θ is ample, this contradicts property (d) of heights. Q. E. D.

By the Mordell-Weil theorem, J_k is a finitely generated abelian group. In particular

$$X = J_k \otimes \boldsymbol{R}$$

is a finite-dimensional real vector space. Moreover, h_θ makes it into a Euclidean space: we will abbreviate the norm $h_\theta(x,y)$ to $\langle x, y\rangle$. The inner product $\langle x, y\rangle$ can be used to compute other heights too:

PROPOSITION 2. *Let $\eta \in Pic(C)$ be a divisor class of degree 0. Then there is a unique point $\tilde{\eta} \in J_k$ such that η equals the restriction of δ_1 to $C \times \{\tilde{\eta}\}$, and*

$$\langle \phi x, \tilde{\eta}\rangle = -h_\eta(x) + O(x), \text{ all } x \in C_k.$$

Proof. The first assertion is part of the definition of the jacobian J of C. The second is an immediate consequence of (b), § 2:

$$\begin{aligned} \langle \phi x, \tilde{\eta}\rangle &= h_\theta(\phi x, \tilde{\eta}) \\ &= h_{(\phi \times 1_J)^*\theta}(x, \tilde{\eta}) + O(x) \\ &= -h_{\delta_1}(x, \tilde{\eta}) + O(x) \\ &= -h_\eta(x) + O(x). \end{aligned}$$

Q. E. D.

PROPOSITION 3. $\langle \phi x, \phi y \rangle = -h_{\Delta^*}(x,y) + O(x,y)$.

Proof. This follows from (c), § 2. Q. E. D.

COROLLARY 1. *There is a constant K such that for $x, y \in C_k$, $x \neq y$,*

$$\langle \phi x, \phi y \rangle \leqq h_{x_0}(x) + h_{x_0}(y) + K.$$

Proof. Recall that $\Delta^* = \Delta - (x_0) \times C - C \times (x_0)$. Apply property (c), § 1 of heights to $h_\Delta(x,y)$; note that the divisor $(x_0) \times C$ (resp. $C \times (x_0)$) is of the form $p_1^*(x)$ (resp. $p_2^*(x_0)$) ; hence $h_{(x_0)\times C}(x,y)$ equals $h_{x_0}(x)$ to within a bounded function and $h_{C\times(x_0)}(x,y)$ equals $h_{x_0}(y)$ to within a bounded function. Q. E. D.

COROLLARY 2. *There is a divisor class $\kappa \in Pic(C)$ of degree 0 such that for $x \in C_k$,*

$$\langle \phi x, \phi x \rangle = 2g h_{x_0}(x) + h_\kappa(x) + O(x).$$

Proof. The self-intersection number (Δ^2) of the diagonal on $C \times C$ is well-known to be $2 - 2g$. Therefore the divisor class on Δ obtained by restricting the class of Δ^* has degree $-2g$. Let

$$f\colon C \to C \times C$$

be the diagonal map. Then there is a divisor class $\kappa \in Pic(C)$ of degree 0 such that

$$f^*(\Delta^*) = -(2g x_0 + \kappa).$$

Therefore

$$\begin{aligned} \langle \phi x, \phi x \rangle &= -h_{\Delta^*}(f(x)) + O(x) \\ &= -h_{f^*(\Delta^*)}(x) + O(x) \\ &= 2g h_{x_0}(x) + h_\kappa(x) + O(x). \end{aligned}$$

Q. E. D.

Putting Proposition 2 and Corollary 1 and 2 together, we obtain the basic estimate:

There is a constant K, and an element $\tilde{\kappa} \in J_k$ such that if $x, y \in C_k$, $x \neq y$, then

$$\langle \phi x, \phi y \rangle \leqq 1/2g \{ \langle \phi x, \phi x \rangle + \langle \phi x, \tilde{\kappa} \rangle + \langle \phi y, \phi y \rangle + \langle \phi y, \tilde{\kappa} \rangle \} + K.$$

4. A packing argument. From here on, we have only to make some elementary observations about Euclidean geometry. First of all, define a new map:

$$C_k \xrightarrow{\psi} X$$

via $\psi(x) = \phi(x) + \frac{\tilde{\kappa}}{2g-2}$. One checks easily that ψ has the property:

$$\begin{cases} \text{There is a constant } K_2 \text{ such that if } x, y \in C_k,\ x \neq y, \text{ then} \\ \langle \psi x, \psi y \rangle \leqq 1/g \left[\frac{\langle \psi x, \psi x \rangle + \langle \psi y, \psi y \rangle}{2} \right] + K_2. \end{cases}$$

Let $\| z \| = \sqrt{\langle z, z \rangle}$, let $f(s) = 1/2(s + 1/s)$, and let

$$\cos(u, v) = \langle u, v \rangle / \| u \| \cdot \| v \|$$

be the cosine of the angle between points $u, v \in X$ in the given norm. Then we can rewrite the above formula as:

$$\cos(\psi x, \psi y) \leqq \frac{1}{g} f\left(\frac{\| \psi x \|}{\| \psi y \|} \right) + \frac{K_2}{\| \psi x \| \cdot \| \psi y \|}$$

Now arrange the countable set of points C_k in a sequence so that

$$\| \psi x_1 \| \leqq \| \psi x_2 \| \leqq \cdots.$$

Note that as $\| \psi x \| \sim \sqrt{2g h_{x_0}(x)}$, (Cor. 2, § 3), it follows that $\| \psi x_i \| \to +\infty$ as $i \to \infty$. The following "packing" lemma is well-known:

LEMMA. *There is an integer N such that if $A_1, \cdots, A_N$ are any non-zero elements of X, then for some pair of integers $1 \leqq i, j \leqq N$,*

$$\cos(A_i, A_j) \geqq \tfrac{2}{3}.$$

COROLLARY. *If $g \geqq 2$ and $\| \psi x_n \| > \sqrt{12 K_2}$, then $\| \psi x_{n+N} \| \geqq \frac{5}{3} \| \psi x_n \|$.*

Proof. If not, whenever $n \leqq i \leqq j \leqq n + N$ then $1 \leqq \| \psi x_j \| / \| \psi x_i \| < \frac{5}{3}$. Hence

$$1 \leqq f\left(\frac{\| \psi x_j \|}{\| \psi x_i \|} \right) < 7/6,$$

and

$$\cos(\psi x_j, \psi x_i) < 7/6g + \frac{K_2}{\| \psi x_i \| \cdot \| \psi x_j \|} < \tfrac{2}{3}.$$

This contradicts the lemma. Q. E. D.

COROLLARY. *If $g \geqq 2$, then there are real constants a and b, $a > 0$, such that*

$$\| \psi x_n \| \geqq e^{an+b}.$$

It is now easy to argue backwards and show that $\| \phi x_n \|$, and $ht_{x_0}(x_n)$, and finally $ht_\delta(x_n)$—for any $\delta \in Pic(C)$ of positive degree—also increase exponentially. This will be left to the reader.

REFERENCES.

I. On the theory of heights:

[1] A. Weil, "Arithmetic on algebraic varieties," *Annals of Mathematics*, vol. 53 (1951), p. 412.

[2] S. Lang, "Les formes bilinéaires de Néron et Tate," *Séminaire Bourbaki*, Expose 274 (1964).

[3] A. Néron, "Hauteurs sur les variétés abéliennes" (to appear).

[4] S. Lang, *Diophantine Geometry*, Interscience-Wiley, N. Y., 1962.

[4½] J. Manin, "The Tate height of points on an abelian variety, its variants and applications," *Izvestia Academii Nauk*, vol. 28 (1964), p. 1363.

II. On the theory of Picard schemes and abelian varieties:

[5] S. Lang, *Abelian varieties*, Interscience-Wiley, N. Y., 1959.

[6] A. Grothendieck, *Fondéments de la géométrie algébrique*, Collected Bourbaki talks, Paris, 1962.

[7] A. Weil, *Variétés Abéliennes et Courbes Algébriques*, Hermann & Cie, Paris, 1948.